COVID-CBR: A Deep Learning Architecture Featuring Case-Based Reasoning for Classification of COVID-19 from Chest X-Ray Images

Xiaohong W. Gao
Department of Computer Science
Middlesex University
London
UK
x.gao@mdx.ac.uk

Alice Gao
A&E Department
Newham University Hospital
Barts Health NHS Trust
London
UK
alicegao@doctors.org.uk

*Abstract— **Background and Objectives**: This study aims to assist rapid accurate diagnosis of COVID-19 based on chest x-ray (CXR) images to provide supplementary information, leading to screening program for early detection of COVID-19 based on CXR images by developing an interpretable, robust and performant AI system. **Methods**: A case-based reasoning approach built upon autoencoder deep learning architecture is applied to classify COVID-19 from other non-COVID-19 as well as normal subjects from chest x-ray images. The system integrates the interpretation and decision-making together by producing a set of profiles that in appearance resemble the training samples and hence explain the outcome of classifications. Three classes are studied, which are COVID-19 (n=250), other non-COVID-19 diseases (NCD) (n=384), including TB and ARDS, and normal (n=327). **Results**: This COVID-CBR system sustains the average sensitivity and specificity of 93.1±3.58% and 96.1±4.10% respectively for classification of these three classes. In comparison with the current state of the art, including COVID-Net, VGG-16 and other explainable AI systems, the developed COVID-CBR system appears to perform similar or better when classifying multi-class categories. **Conclusion**: This paper presents a case-based reasoning deep learning system for detection of COVID-19 from chest x-ray images. Comparison with several state of the art systems is conducted. Although the improvement tends to be marginal, especially for VGG-16, the novelty of this work manifests its interpretable feature building upon case-based reasoning, leading to revealing this viral insight and hence ascertaining more effective treatment and drugs while maintaining being transparent. Furthermore, different from several other current explainable networks that highlight key regions or the points of an input that activate the network, i.e. heat maps, this work is constructed upon whole training images, i.e. case-based, whereby each training image belongs to one of the case clusters.*

Keywords— case-based reasoning (CBR), chest x-rays (CXR), classification, COVID-19, deep learning

I. INTRODUCTION

This study presents an interpretable deep learning system built upon case-based reasoning, to classify COVID-19 from normal and other respiratory diseases from chest x-ray images, complimenting the current laboratory methods for testing COVID-19, leading to improved diagnostic accuracy.

1.1 Covid-19

The novel coronavirus 2019 (COVID-19), a classified pandemic by World Health Organisation (WHO) in March 2020 [1] and commenced in late 2019 [2, 3], has shattered the world by infecting more than 120 million people and caused over 2.6 million deaths [4].

Officially known as SARS-CoV-2, COVID-19 is a strain of coronavirus that remains contiguous and can transmit from person to person. The clinical symptoms of Covid-19 can range from a mild common cold-like symptoms, to a severe viral pneumonia leading to acute respiratory distress syndrome (ARDS), advancing to potentially fatal. At present, the presence of COVID-19 in respiratory specimens is detected by next generation sequencing or real-time reverse transcription polymerase chain reaction (RT-PCR) methods, the gold standard of a laboratory technique combining reverse transcription of Ribonucleic acid (RNA) into Deoxyribonucleic acid (DNA) and amplification of specific DNA targets. Another type of test is to take on antibody samples, by which a blood sample is collected to check for antibodies.

Because there remains many unknowns regarding to Covid-19, every test has its limitations on [5] conclusive diagnosis, usually with low specificities especially at the onset of COVID-19. Hence, it is important to provide additional information to corroborate diagnosis as any false negative diagnosis can instigate potential spreading risk to other people unknowingly.

As confirmed cases continues to increase considerably globally, timely detection of the disease not only can provide supportive care required by patients but also can prevent further spread of the virus. Consequently, effective screening infected patients using CXR accounts for a critical step in this fight against COVID-19 as well as to circumvent temporary shortage of RT-PCR kits in confirming COVID-19 infection.

This study works on the detection of COVID-19 from chest x-ray images employing a case-based reasoning approach, aiming at providing complimentary information to improve the accuracy of diagnosis while minimising invasiveness while maintaining the developed deep learning system transparent and interpretable. With regard to medical images for diagnosis of COVID-19, Computed Tomography (CT) and Chest X-ray (CXR) represent the most common imaging tools. While x-ray machines remain one of the first imaging tools in clinics with advantageous features of being portable, non-invasive, fast image acquisition, economic and much less exposing to radiation in comparison with Computerised Tomography (CT), 3D CT lung images provide high resolution and detailed information. On a CXR, the most common features of COVID-19 appear to be

bilateral infiltrated with peripheral opacities and patchy consolidation [6, 7] whereas bilateral ground glass opacities retain a key finding on CT [8, 9].

The challenge facing detecting COVID-19 from CXR images is that when the disease is at its early onset, the characteristic patterns present less obvious to the human eyes [10]. However, as more cases arrive, more COVID-19-specifc patterns will be determined. Progress on detection of COVID-19 from CXR has been made significantly recently with a plethora of research work being published, in particular by means of deep learning techniques.

A. Deep learning networks for detection of COVID-19

From 2020, a large number of work has been conducted applying AI techniques, in particular, deep learning, to predict COVID-19 virus and have demonstrated significant performance. For example, for 3D CT images, attention-based deep learning networks have shown effectiveness in classifying COVID-19 from normal subjects [11, 12]. In relation to chest x-ray images, patch-based CNN is applied to study chest x-ray images [13] and to differentiate discriminatory features of COVID-19. Specifically, COVID-Net [14], one of the pioneer studies, classifies COVID-19 from normal and other pneumonia diseases through the application of a tailored deep learning network. COVID-Net starts with a pre-trained model of COVID-Net-S and endorses the design pattern of projection-expansion-projection coupled with features of high architectural diversity and selective long-range connectivity. To overcome the shortage of datasets, a number of researchers [15] apply generative adversarial neural network (GAN) to augment data first and subsequently to classify COVID-19.

With regard to chest x-ray images, significant work has also been conducted on. For example, the work in [16] evaluates five pre-trained CNN backbone of ResNet series of models and concludes that ResNet50 present the best performance. In the studies in [17, 18], VGG series of networks as well as ResNet models are evaluated including MobileNetV2, AlexNet, VGG16, VGG19 and GoogleNet. Understandably, to a large extend, the performance of each network is dependent on the dataset properties, including, size, disease types and classification categories, e.g. binary classification (e.g. COVID-19 vs non-COVID-19) usually presents better performance than multi-class classification (e.g. COVID-19, Normal and other pulmonary diseases).

While deep learning techniques have achieved state of the art results, giving rise to becoming an integral and indispensable approach in assisting people to process big data in the current digital era, its transparency and interpretability become increasingly important, especially in clinical practices [19]. Hence, progress towards the development of explainable deep learning networks have been made more recently [20, 21] for prediction of COVID-19 applying heat maps [22] to visualise the activation maps. However, the networks that generate these heat maps are usually not part of the original training architectures and apply different sets of parameters to start, to train and to conclude, which leads challenges to interpret together with the original input images.

Therefore, in this study, a deep-learning network is developed with a feature of case-based reasoning (CBR), coined as COVID-CBR, to assist the diagnosis of COVID-19 from CXR, aiming at reliably identifying infected patients with a low rate of both false negatives and false positives while elucidating the reason behind detection results. Specifically, different from the other existing deep-learning based computer systems, the decisions made by this COVID-CBR will make a step closer towards being transparent and interpretable to clinicians.

B. Case-based reasoning deep learning neural networks

In the evidence-based medical domain, cases are the most specialised form of knowledge representation, consisting of both general understanding and clinicians' experiences, taking into considerations of differences between their current patient and typical or known exceptional cases [23]. This process is known as case-based reasoning (CBR) whereas the generalised cases are termed as prototypes [24] or profiles.

At present, the interpretation of a deep learning network is realised in two steps. Firstly, a neural network is designed to achieve state of the art results. Then, its interpretability is analysed after the training by setting up a separate model to decipher the achieved results, e.g. studying saliency map features. This however gives rise to the credibility of attained explanations as interpretability analysis derives from a different modelling process with priors that are not part of the training from the original networks [25]. To ensure the interpretation of a network is meaningful, understandable and credible, much research has since been conducted with a focus on the visualisation of parts of images that most strongly activate a given feature map [26]. More recently, progress has been made to allow case-based interpretation through prototyping [27]. Rather than imposing an additional structure on feature maps, the prototype-based approach introduces a special prototype layer for explanation of decision making within the same training network.

Inspired by the work in [28] and autoencoder [29] architecture, this study builds an enhanced case-based reasoning deep learning network to classify COVID-19 from normal and other lung diseases on chest x-ray images. This network models a profile layer comprising a list of profiles whereby each profile resembles observations in one of classes in visual appearance. Hence this set of profiles learns toward being a representative of the whole training set, which intends to be significant beneficial as more data will determine more COVID-19 specific features.
.

II. METHODS

A. Data collection

The datasets in this study are collected from public resources and consist of 961 posteroanterior (PA) view (back projection) chest radiograph images, making up COVID-19 (n=250), other lung diseases (n=384), and normal subjects (n=327). The public available sources include *COVID-19 Image Data Collection* [30], *Figure1 COVID-19 Chest X-ray Dataset Initiative* [31], *Databiology* [32], *Italian Society of Medical and Interventional Radiology* [33] and *Radiopeadia* [34], at which images are constantly updated. In addition, two public x-ray datasets for both Tuberculosis (TB) and normal subjects are collected [35]. These data are grouped into three clusters, which are COVID-19, non-

COVID-19 diseases (NCD) (including TB, ARDS, SARS, E.Coli, Pneumocystics, Streptococcus, Chlamydophila, Legionella, Klebsiella) and normal as exemplified in Fig. 1 where arrows point to the diseased regions.
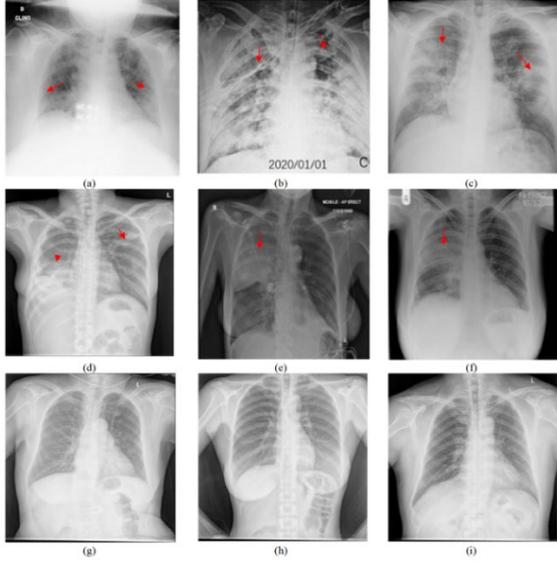
Fig 1. Examples of chest x-rays of COVID-19 ((a) to (c)) in comparison with the other pulmonary diseases ((d): TB, (e) ARDS, (f) SARS) and normal ((g) to (i))). Arrows refer to diseased regions.

The top row of Figure 1 comprises COVID-19 disease ((a) to (c)) whereas middle row the other pulmonary diseases ((d) TB, (e) ARDS, (f) SARS). Figure 1(g) to Figure 1(i) display CXR images from normal subjects. For COVID-19, the visual patterns include bilateral infiltrated with peripheral opacities (Figure 1(a)), patchy consolidation (Figure 1(b)), and bilateral ground glass opacities (Figure 1(c)). For TB (Figure 1(d)), consolidation (arrowhead) and left paratracheal stripe (arrow) are present. With regard to ARDS (Figure 1(e)), airway stenosis is shown with patchy consolidation. Figure 1(f) illustrates an example of SARS presenting airspace opacities in the middle or lower zones.. While these non-COVID-19 lung diseases (Figures 1(d) to (f)) share several visual features, e.g. SARS may show ARDS types with confluent consolidation, they are different from COVID-19 that exhibits bilateral patterns.

In this collection, due to data coming from different resources, such as publications of various journals, the resolutions of images vary between 150×150 and 4200×3400 pixels. Hence these images are firstly normalised to 512×512 and then converted to MNIST [37] dataset format for expedition of training process. The ratio between training and testing data retains 90:10.

B. Deep learning network with case-based reasoning for classification of COVID-19

As illustrated in Fig. 2, this proposed case-based reasoning architecture incorporates four components: encoder, decoder, classifier and reasoning profiles. The network is analogous to an autoencoder architecture, where the profiles, $(p_1, p_2, \ldots p_m)$, as well as the classifier are in the latent space. These profiles are expected to provide common features hence the explanation of the decision making towards classification by producing similar images in appearance to one of classes.

Hence, when a test image is loaded to the trained model, the model calculates the overall distance between this test image and each of the profile images and delivers the final classification result.
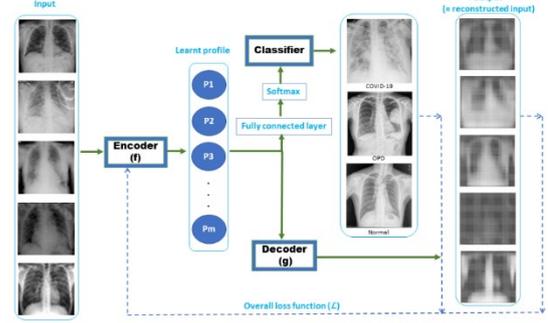
Fig. 2. The proposed case-based reasoning profiling network that interprets the classification.

The function of encoder aims to reduce the dimensionality of the input (as well as noise) and to learn the weights (W) of transformation from input, leading to the final prediction of classes using Eq. (1), whereas the profile layers (P) in between generates the profile units ($p_i, i = 1, \ldots, m$) that resemble in appearance one of the K classes. In our case, $K = 3$, which symbolizes 'COVID-19', 'NCD' for other lung diseases (e.g. pneumonia, TB, ARDS, etc.), and 'normal'.

$$P = f(X) = f'(WX + B) \quad (1)$$

In Eq.(1), the input $X = (x_1, x_2, x_3, \ldots, x_n)^T$ indicates n samples with each image ($x_i, i = 1, \ldots, n$) having a size of 512×512 and produces an set of profiles $P = (p_1, p_2, p_3, \ldots, p_m)^T$. In addition, B represents the bias that is initiated randomly during the training. The profile number (m) (i.e. different visual patterns of training dataset) is pre-defined and can have the size of class numbers (K) or more (in this study, m=30). W stands for the weight matrixes in the encoder that are to be determined in the training process.

Specifically, f' refers to the transformation from a range convolution layers to profiles in an encoder as illustrated in Figure 2. In this study, the magnitude of m varying from 3 to 45 is investigated. It has been found that more profiles do not necessarily lead to more accurate results as over-sized profiles appear to lead many being redundant by presenting near blank features.

This profile layer measures the squared distance between encoded input z (Eq. (2)) and each of the profile vectors as formatted in Eq. (3).

$$z = [f(x_1), f(x_2), f(x_3), \ldots, f(x_n)] \quad (2)$$
$$P(z) = [\sum(z - p_1)^2, \sum(z - p_2)^2, \ldots, \sum(z - p_m)^2]^T \quad (3)$$

After the profile layer, a fully connected layer and a classification layer follow to compute weighted sums of these distances $W_P(P(z))$, where W_P is the $K \times m$ weight matrix and will be learnt by way of training as illustrated in Figure 2. These weighted sums are then normalized by the *Softmax* layer to produce a probability distribution over the K classes. Hence, the distribution of probability of a test image that belongs to each class is calculated in the *Softmax* layer resulting in a form of a vector with K elements, where the k_{th}

$(k = 1,.., K)$ component of the output of the *Softmax* layer is defined by Eq. (4).

$$k_{th} = S_{softmax}(v_k) = \frac{\exp(v_k)}{\sum_{i=1}^{K} \exp(v_i)} \qquad (4)$$

In Eq. (4), v_k is the k_{th} component of the vector $V = W_P(P(z)) = (v_1, \ldots, v_k)$.

During the prediction, the neural network architecture depicted in Figure 2 delivers the class label that has the highest probability among the K vector derived from Eq. (4).

In Fig. 2, the *Decoder* reconstructs back the input $x \in X$, based on the profiles, i.e. from $m \times 1$ profile units to construct 512×512 image using a function g expressed in Eq. (5), which decodes the encoded feature vectors in $x, x \in X$.

$$x_- = g(x) \qquad (5)$$

Hence, the multi-task loss function ($\mathcal{L}$) for the network of Figure 2 is formulated in Eq. (6) by combining the loss of classification, decoding and two interpretability regularisation measures.

$$\mathcal{L} = \lambda_1 \mathcal{L}_{classification} + \lambda_2 \mathcal{L}_{decoder} + \mathcal{L}_{interpreter-1} + \lambda_4 \mathcal{L}_{interpreter-2} \qquad (6)$$

where λ_1 to λ_4 are the real valued hyperparameters and applied to adjust the ratios between those four terms.

The classification loss applies the standard cross-entropy function as given in Eq. (7).

$$\mathcal{L}_{classification} = \frac{1}{n} \sum_{i}^{n} (y_i * \log(\hat{y}_i)) \qquad (7)$$

where n is the total number of data samples with y_i referring to the i_{th} sample label and $\hat{y}_i$ the predicted label.

Furthermore, as expressed in Eq. (8), the loss function for the reconstruction of decoding is quantified using mean squared errors (MSE).

$$\mathcal{L}_{decoder} = \frac{1}{n} \sum (X - X_-)^2 \qquad (8)$$

Similar to the work described in [28], the two interpretability measures are calculated using Eqs. (9) and (10), which are established to safeguard respectively the distance of each profile to be as close as possible to at least one of the training samples in the latent space, and the distance of each encoded training sample to be as close to one of the profiles as possible.

$$\mathcal{L}_{interpreter-1} = \frac{1}{m} \sum_{j=1}^{m} \min((p_j - f(x_1))^2, (p_j - f(x_n))^2) \qquad (9)$$

$$\mathcal{L}_{interpreter-2} = \frac{1}{n} \sum_{i=1}^{n} \min((p_1 - f(x_i))^2, (p_m - f(x_i))^2) \qquad (10)$$

In this way, $\mathcal{L}_{interpreter-1}$ will propel the profile vectors to have meaningful decoding in the pixel space, whereas $\mathcal{L}_{interpreter-2}$ will cluster the training samples closely around profiles in the latent space. Therefore, these two measures are anticipated to usher profiles to training samples in a tight closeness in visual appearance.

III. RESULTS AND ANALYSIS

The implementation takes place applying Python with TensorFlow and Keras libraries. In this study, the values of λ_1 to λ_4 are set to 0.85, 0.05, 0.05, and 0.05 respectively, allotting highest weight to classification, which appears to deliver optimal results in comparison with other combinations. Similar to conventional CNN, the encoding

process is composed of six convolutional layers with each one having a filter size 3×3.

After training for 3000 epochs, for classification of three classes of ''COVID-19, 'NCD, and 'normal', the proposed model (Figure 2) achieved averaged sensitivity of 93.1 ± 3.48% and specificity of 96.1 ± 4.10% when the profile numbers (m) are set to 30. The testing result is based on 5 runs with each run taking 20 test samples.

Fig. 3 demonstrates the performance of decoding to reproduce twenty training samples (top row) (randomly selected from test set) using trained profiles. Visually, the regenerated samples (bottom row) appear to be close to the original images (top row), indicating the profiles tend to be representative of the training samples.

(a)

(b)
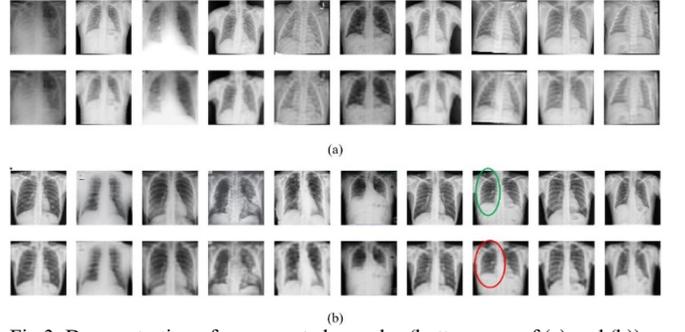
Fig.3. Demonstration of re-generated samples (bottom row of (a) and (b)) using the trained model of profiles from the original images (top row of (a) and (b)).

While most of the samples are re-constructed well in appearance, all the re-generated images appear to be less sharp than the original ones. For example, the red circle in Figure 3(b) exhibits less clear horizontal lines (e.g. ribs) than the above original image marked in green circle. However, when it comes to the classification of COVID-19, this smoothing feature tends to be insignificant in comparison with other deep learning models as given in Table 1.

In Fig. 4, the thirty profiles are presented, which are trained to be representative of three classes (i.e. 'COVID-19, 'NCD', and 'normal') of the training samples. While in appearance, those profiles depict less resemblance to a proper CXR image, they intend to encompass distinguishing features between classes.
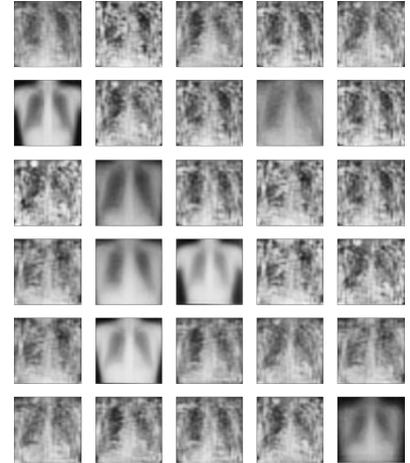
Fig. 4. The thirty profiles that represent training samples of three classes, i.e., 'COVID-19, 'NCD', and 'normal', where * refers to centroid of each class.

Table 1 lists the classification result for this developed COVID-CBR system together with current state of the art deep learning systems, including VGG-16, COVID-Net and COVID-Net-Shuffle. All these systems are trinaing using the same set of data. VGG-16 denotes a transfer learning classification network containing nine convolutional layers built upon a pre-trained model VGG16 that comprises 16 layers, whereas COVID-Net-Shuffle signifies the architecture built upon ShuffleNet [38] to enhance the calculation efficiency by balancing the performance between the calculation speed and accuracy.

TABLE 1. COMPARISON OF CLASSIFICATION RESULTS BY COVID-CBR, VGG-16, COVID-NET-SHUFFLE AND COVID-NET SYSTEMS USING THE SAME SET OF DATA. SPE=SPECFICITY; SEN=SENSITIVITY.

		COVID-19	NCD	Normal	Average (%)
COVID-CBR	Spe	98.6±0.57	87.1±4.17	93.5±6.01	**93.1±3.58**
	Sen	96.5±3.19	95.6±3.94	96.3±5.17	**96.1±4.10**
VGG-16	Spe	97.6±0.94	88.3±7.86	93.5±1.91	93.1±3.57
	Sen	96.6±1.27	95.3±1.11	94.7±3.75	95.5±2.04
COVID-Net-Shuffle	Spe	92.6±2.0	84.4±5.78	89.1±2.94	88.7±3.08
	Sen	90.7±3.65	89.4±4.81	93.9±1.91	91.3±2.66
COVID-Net	Spe	90.2±7.32	80.9±2.13	90.2±4.16	87.1±4.53
	Sen	93.1±2.46	91.1±1.62	95.9±1.98	93.4±2.02

Table 1 depicts that this developed COVID-CBR deep learning system performs the best and can predict COVID-19 with 98.6% sensitivity and 96.5% specificity whereas 93.5% sensitivity and 96.3% specificity are achieved for the normal subjects. For the other pulmonary diseases, i.e. class of 'NCD', while the specificity tends to be higher at 95.6%, the sensitivity retains only 87.1%, implying that while the other two classes (e.g. COVID-19 and Normal) are less likely to be classified into this 'NCD' class, this class is much easier to be categorised into being either COVID-19 or Normal. Overall, the average sensitivity and specificity for the three classes retain 93.1 ± 3.58% and 96.1 ± 4.10% respectively, which are attained base on 5 runs of testing with each run accommodating 20 test samples.

The performance based on transfer learning built upon VGG16 appear to be very similar to the results by COVID-CBR with the average of 93.1% and 95.5% sensitivity and specificity respectively with slightly better performance on NCD class with 88.3% sensitivity (87.1% for COVID-CBR). For COVID-Net-Shuffle, the average sensitivity and selectivity are 88.7% and 91.3% whereas 86.1% and 92.65% are realised for COVID-Net.

Overall, all approaches listed in Table 1 exhibit higher specificity for prediction of COVID-19 with 96.5%, 96.6%, 90.7% and 95.6% respectively for COVID-CBR, VGG-16, COVID-Net-Shuffle and COVID-Net, suggesting that the patterns of COVID-19 on x-rays present appreciably unique characteristics comparing with the other two classes, from which those two classes are less likely to be classified into COVID-19.

Comparison with the other multi-class classification systems is provided in Table 2 where the results are directly obtained from their publications with different number of datasets (in brackets).

TABLE 2. COMPARISON OF SENSITIVITY (ACCURACY) WITH PUBLISHED WORK FOR MULTI-CLASS DETECTION AND WITH EXPLAINABLE NETWORKS BASED ON CHEST X-RAY IMAGES. THE NUMBERS IN BRACKETS REFER TO DATA NUMBERS.

Architecture	COVID-19 (Spe (%))	Non-COVID-19 (Spe (%))	Normal (Spe (%))	AVG (Total)
VGG-19 [14] *(Data number)*	58.7 *(358)*	90.0 *(5,538)*	98.0 *(8,066)*	82.2 *(13,962)*
ResNet-50 [14] *(Data number)*	83 *(358)*	92 *(5,538)*	97 *(8,066)*	90.6 *(13,962)*
COVID-Net [14] *(Data number)*	91.0 *(358)*	94 *(5,538)*	95 *(8,066)*	93.3 *(13,962)*
Transfer Learning [29] *(Data number)*	96 *Stage 2 (1,150)*	99.1 *Stage 1 (3,303) Stage 2 (2,753)*	95.7 *Stage 1 (3,520)*	**96.9** *(6,823) (3,903)*
COVIDx-YOLO [28] *(Data number)*	94.1 *(125)*	87.0 *(500)*	88.0 *(500)*	87.02 *(1,125)*
COVID-CBR *(Data number)*	**98.6** *(250)*	87.1 *(384)*	93.5 *(327)*	93.1 *(961)*

One of the challenges facing detection of COVID-19 is multi-class predictions, containing COVID-19, other pulmonary diseases and normal subjects. Hence comparison with the other systems for multi-class detection, also for explainable networks is provided in Table 2. Each study applies varying datasets. Understandably, larger datasets usually produces better results. For example, in Table 2, COVID-Net (n=13,962) has an averaged specificity of 93.3% in Table 2 and 87.1% in Table 1 with a total number of data 961. While transfer learning appears to perform the best with 96.9% specificity, in essence, at each stage, the network conducts binary classification. For prediction of COVID-19, our method of COVID-CBR out performs the others with 98.6% specificity.

IV. DISCUSSION AND CONCLUSION

This work aims at establishing a reasoning system based on experts' knowledge to predict COVID-19 disease from chest radiographs. This case-based reasoning system, COVID-CBR, is deployed applying deep learning techniques and appears to have achieved better results in multi-class classification for COVID-19, NCD and normal categories.

While most of the current work focuses on two or three classes, detailed classification is conducted by Oh et al [13] [39], comprising categories of normal, tuberculosis, bacteria, viral and COVID-19. Due to limited dataset (n=502, viral=20), they segment lung regions first into patches before performing statistical analysis of characteristics of each class. The size of patches appears to affect the classification results, with the best sensitivity being 96.4%. In addition, the authors also apply Grad-CAM approach to generate heat maps for representing the system activations.

This COVID-CBR network, however, utilises case-based reasoning instead of extractive reasoning by explaining its predictions based on similarities between observations and profile cases, rather than highlighting the most relevant parts of the input that in many cases are not determined.

Significantly, this COVID-CBR is probably the first system that employs case-based reasoning to predict COVID-

19 and has demonstrated to be able to provide promising results. Furthermore, the predicting process manifests near-transparency features by building upon existing known cases, in the form of profiles, which can potentially be applied to assist clinicians' decision making while performing screening of this deadly devastating COVID-19 virus.

Since these conclusions are based upon a small sample size (n=961), further study will collect more data sets and may also take 3D CT images into considerations. It is expected that more data will foster more comprehensive profiles.

While more profiles (e.g. case representative) may cover variations of training samples, exceeding numbers may not necessarily maintain better accuracy with many appearing to be redundant.

Since training takes place using the conventional 6-layer CNN structure (plus one fully connection layer) without transfer learning, small sample size will make considerable impact to the training process. Again, data enlargement will become crucial in the future to collect more cases that can then in turn be reasoned into more profiles.

In conclusion, this work develops a case-based reasoning system for classification of COVID-19 from chest x-ray images and achieves state of the art results, especially for predicting COVID-19 class, contributing to the development of robust, transparent and performant AI systems for medical applications.

References

[1] WHO, Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020, Who.int. 2020. https://www.who.int . Retrieved in April 2020.

[2] Zhu N, Zhang D, Wang W, Li, X, et al, A Novel Coronavirus from Patients with Pneumonia in China, 2019 , New England Journal of Medicine, 382(8):727-733, 2020.

[3] Perlman, S., Another Decade, Another Coronavirus. (2020) New England Journal of Medicine, 382(8): 760:762, 2020.

[4] Resource Centre, Johns Hopkins University and Medicine, https://coronavirus.jhu.edu/.

[5] COVID-19 Guidance, Medicines & Healthcare products Regulatory Agency, https://www.gov.uk/ . Retrieved in October 2020.

[6] Chen N, Zhou M, Dong X, et al., Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study, The Lancet, 395 (10332):507:513, 2020.

[7] Huang, C., Wang, Y., Li, X., et al, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet, 395 (10223):497-506, 2020

[8] Raptis C, Hammer M, Short R, et al. Chest CT and Coronavirus Disease (COVID-19): A Critical Review of the Literature to Date, American Journal of Roentgenology, 1-4. 10.2214/AJR.20.23202, 2020.

[9] Kanne J, Little B, Chung J, Elicker B, Ketai L, Essentials for Radiologists on COVID-19: An Update—Radiology Scientific Expert Panel, Radiology, 296(2):E113-E114, 2020.

[10] Ng M, Lee E, Yang J, et al, Imaging Profile of the COVID-19 Infection: Radiologic Findings and Literature Review, Radiology: Cardiothoracic Imaging, 2:1, 2020.

[11] Wang J, et al., Prior-attention residual learning for more discriminative COVID-19 screening in CT images, *IEEE Trans. Med. Imag.*, early access, May 15, 2020.

[12] Ouyang X, *et al.*, Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia, *IEEE Trans. Med. Imag.*, early access, May 18, 2020.

[13] Oh Y, Park S, Ye JC, Deep learning COVID-19 features on CXR using limited training data sets, *IEEE Trans. Med. Imag.*, early access, May 8, 2020.

[14] Wang L, Lin Z, Wong A, COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest Radiography Images}, arXiv.2003.09871, 2020

[15] Waheed A, Goyal M., Gupta D, Khanna A, Al-Turjman F, Pinheiro PR, , CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection, IEEE Access, May 28, 2020.

[16] Wexler R., When a computer program keeps you in jail: How computers are harming criminal justice, *New York Times*, 2017.

[17] Song Y, Zheng S, Li L, et al, Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images, medRxiv, 2020.

[18] Narin A, Kaya C, Pamuk Z, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, arXiv preprint arXiv:2003.10849, 2020.

[19] Hemdan EED, Shouman MA, Karar ME, COVIDx-Net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images, arXiv preprint arXiv:2003.11055, 2020 .

[20] Sethy PK, Behera SK, Detection of coronavirus disease (covid-19) based on deep features, Preprints, 2020030300 (2020), p2020, 2020.

[21] Schmidt R, Waligora T, Vorobieva O, Prototypes and Case-Based Reasoning for Medical Applications, Studies in Computational Intelligence (SCI), 73:285–317, 2008.

[22] Ozturk T, Talo M, Yildirim EA, *et al*., Automated detection of covid-19 cases using deep neural networks with x-ray images, Computers in Biology and Medicine, p. 103792, 2020.

[23] Brunese L, Mercaldo F, Reginelli A, Santone A, Explainable Deep Learning for Pulmonary Disease and Coronavirus COVID-19 Detection from X-rays, Computer Methods and Programs in Biomedicine, 196:105608, 2020.

[24] Selvaraju RR, Cogswell M, Das A, *et al.* , Grad–cam: Visual explanations from deep networks via gradient-based localization, in: Proceedings of the IEEE international conference on computer vision, 2017, pp. 618–626, 2017.

[25] Schmidt R, Waligora T, Vorobieva O, Prototypes and Case-Based Reasoning for Medical Applications, Studies in Computational Intelligence (SCI), 73:285–317, 2008.

[26] Swanson DB, Feltovich PJ, Johnson PE, Psychological Analysis of Physician Expertise: Implications for Design of Decision Support Systems. In: Shires DB, Wolf H (eds.): Proceedings of MEDINFO 77, pp161–164, 1977.

[27] Montavon G., Samek W., and Müller K. (2017), Methods for interpreting and understanding deep neural networks, CoRR abs/1706.07979, 2017.

[28] Pinheiro PO, Collobert R, From image level to pixel-level labelling with convolutional networks, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1713–1721, 2015.

[29] Kolodner J, An introduction to case-based reasoning, Artificial Intelligence Review, 6 : 3-34, 1992.

[30] Li O, Liu H, Chen C, Rudin C, Deep learning for case-based reasoning through prototypes: A neural network that explains its predictions, AAAI 2018, 2018.

[31] COVID-chest- x-rays: https://github.com/ieee8023/covid-chestxray-dataset.

[32] Cohen J, Morrison P, Dao L, COVID-19 image data collection. arXiv 2003.11597, 2020.

[33] DataBiology, https://www.lab.databiology.net/ . Accessed in October 2020.

[34] Italian Society of Medical and Interventional Radiology, https://www.sirm.org/en/italian-society-of-medical-and-interventional-radiology/. Accessed in October 2020.

[35] Bell, D, Knipe H, et al, COVID-19 (summary), Radiopaedia, https://radiopaedia.org/articles/covid-19-summary?lang=gb. Retrieved in October, 2020.

[36] Ma N, Zhang X, Zheng H, Sun J, ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design, arXiv:1807.11164, 2018.

[37] MNIST, http://yann.lecun.com/exdb/mnist/. Received in September, 2021.

[38] Ma N, Zhang X, Zheng H, Sun J, ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design, arXiv:1807.11164, 2018.

[39] Ester Zumpano, Antonio Fuduli, Eugenio Vocaturo, Matteo Avolio: Viral pneumonia images classification by Multiple Instance Learning: preliminary results. IDEAS 2021: 292-296